DOCUMENT RESUME

ED 174 621 **SP 014 720**

AUTHOR Martin, Jeanne; And Others

TITLE The Unit of Analysis Problem in Teacher Effectiveness Research.

INSTITUTION Texas Univ., Austin. Research and Development Center for Teacher Education.

SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.
PUB DATE Feb 79
CONTRACT NIE-C-74-0089
GRANT OB-NIE-6-78-0116; OE-NIE-6-0216
NOTE 13p.

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Educational Research; *Effective Teaching; *Evaluation Methods; Primary Education; Research Criteria; *Statistical Analysis; Student Teacher Relationship; *Teacher Evaluation; *Teaching Methods

ABSTRACT

Although teacher effectiveness research has to date been primarily concerned with class level statistics, valuable information regarding a teacher's proficiency can be gleaned from the analysis of within class data. The relationships between teacher behavior and student achievement vary at the differing levels of analysis. Nine variable patterns of teacher behavior yield alternative response on more than three categorical levels: students within reading groups; reading groups within classes; and classes. Cross-group study of these levels, when compared with in-group study, provides information concerning not only the competency of the individual teacher but also the relativity of the individual student's achievement level to that student's social group. A synthesis of the two types of statistical analysis is necessary to obtain a comprehensive view of the dynamism between student-teacher interaction and student achievement. (LH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED174621

The Unit of Analysis Problem in Teacher Effectiveness Research

Jeanne Martin

Linda M. Anderson

Donald J. Veldman

The Research and Development Center for Teacher Education

The University of Texas at Austin

February 1979

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

This study was supported in part by the National Institute of Education under Contract NIE-C-74-0089 and Grants OB-NIE-6-78-0116 and OB-NIE-6-0216, The Research and Development Center for Teacher Education. The opinions expressed herein do not necessarily reflect the position or policy of the National Institute of Education and no official endorsement by that office should be inferred.

SP 014 720

The Unit of Analysis Problem in Teacher Effectiveness Research

Teacher effectiveness studies are concerned with determining the relationship between teacher behaviors occurring during the process of instruction and student learning. Briefly, teacher effectiveness research involves observing and recording events which occur in the classroom over a specified period of time. Student achievement is measured before and after the period of observation. The increase in student achievement which occurs over the observation period is then related to the occurrence of teacher behaviors. The goal of this type of research is to identify teacher behaviors which have been effective in promoting student learning in the classroom.

Teacher effectiveness researchers are frequently confronted with the problem of determining the appropriate unit of analysis to use in analyzing data which have been collected on individual students in intact classrooms. Traditionally, class means have been chosen to analyze this type of data to avoid the methodological problems which arise when student scores are pooled across classes (see Glendening and Porter, 1974). However, when data are analyzed with class means, information about individual students is lost. This is particularly distressing when data are collected at the individual level, when there are few classes involved in the study, and when the researcher desires to study interactions involving student characteristics. Recently, however, Cronbach (1976) has presented an innovative analysis approach which resolves many of these problems; Cronbach advocates analyzing data within classes as well as at the class level. Within class analyses avoid the methodological problems encountered when student scores are pooled across classes and yet provide information about relationships occurring at the student level within classes--information which would otherwise be lost in class level analyses.

The investigation reported in this paper was designed to compare relationships obtained between teacher behaviors and student achievement in within-class analyses with those relationships obtained in between-class analyses. Such a comparison would determine whether within-class analyses are desirable, i.e., provide information beyond that obtained in class mean analyses. It was hypothesized that relationships obtained at the class mean level of analysis would not be the same as relationships obtained within classes and vice versa, thereby verifying the value of both types of analyses.

Methodology

The data for this investigation were taken from the First-grade Reading Group Study (Anderson, Evertson, and Brophy, 1978). In this study, individual interactions between first-grade students and teachers were recorded as the students participated in reading groups within intact (self-contained) classrooms. Therefore, there were three potential analysis units: class means, reading group means, and student scores. It was decided to analyze student scores within reading groups, reading group means within classes, and to also analyze the data with class means. Each of these three types of analysis will address somewhat different questions, since the sources of variance at each level are somewhat different.

There were a total of 15 classes, 41 reading groups, and 282 students. Only 18 out of 473 student-teacher interaction variables recorded in the First-grade Reading Group Study were investigated. However, additional variables will be included in future analyses and a more comprehensive report will be made available concerning specific results for the variables studied. A list of the student-teacher variables included in the present investigation

are given below:

1. method by which teacher selects students to answer questions or to read (ordered turns, preselections, volunteer selections, nonvolunteer selections, and call outs)
2. type of teacher question (reading or nonreading)
3. type of student answer (correct, incorrect, "don't know" response, or no response)
4. correct answers followed by no feedback
5. correct answers followed by emphasis feedback, i.e., the teacher does something to insure the answer just given by a student is heard by other students
6. correct answers followed by teacher praise
7. correct answers followed by the teacher asking the student a new question
8. incorrect answers followed by teacher criticism
9. incorrect answers followed by sustaining feedback, i.e., the teacher requires the student to continue responding
10. incorrect answers followed by terminating feedback, i.e., the teacher ends the student's response opportunity by getting the answer from another student or giving the answer herself.

Three sets of regression models were evaluated for each variable to determine its relationship to achievement. In all models, readiness scores were used as covariables. The first set of models contained class means as the unit of analysis. The second set of models contained reading group means as the unit of analysis with a set of binary vectors indicating class membership; thus, reading group means were analyzed within classes. The third set of models contained individual student scores as the unit of analysis with

a set of binary vectors indicating reading group membership so that student scores were, in effect, analyzed within reading groups.

Results

Results showed that relationships between teacher behaviors and student achievement are generally not the same when different units--students within reading groups, reading groups within classes, and classes--are used in the analyses. For example, at the student-within-reading-group level of analysis, 11 out of 18 variables were significantly related to achievement; at the reading-group-within-class level, three variables were significantly related to achievement; and at the class level four variables were significantly related to achievement.

The greater number of significant relationships at the student-within-reading-group level of analysis may be due, at least partially, to the increase in power at this level, as the degrees of freedom have increased from 11 for class level analyses to 23 for reading-group-within-class analyses to 244 for student-within-reading-group analyses. However, beyond approximately 30 degrees of freedom, the increase in power which accompanies an increase in the degrees of freedom levels off (Myers, 1972). Therefore, a three-fold increase in significant results at the student-within-reading-group level of analyses cannot be wholly attributed to increases in power. Also, if differences in significant relationships at the different levels of analysis were to be attributed solely to power increases, one would expect considerably more significant results at the reading group level of analysis than at the class level. Yet in this study, more significant relationships were obtained at the class level of analysis than at the reading-group-within-class level.

The lack of consistency across different levels of analysis is not surprising when one considers that analyses with different units of analysis answer different questions. Analyses with class means as the unit of analysis determine whether class means on the teacher behavior are related to mean class achievement. Because class means are involved, relationships obtained in these analyses are relationships between absolute (as opposed to relative) amounts of the behavior. In contrast, reading-group-within-class analyses determine whether, within classes, reading group means are related to achievement. Because reading group means are compared within classes, relationships obtained in these analyses are relative. That is, a positive relationship indicates that reading groups which have received a higher amount of the teacher behavior relative to other groups within the same class will achieve more--regardless of the actual amount of the teacher behavior received by that higher group. Similarly, individual-within-reading-group analyses reveal whether student scores within reading groups are related to student achievement. Because student scores are compared within reading groups, relationships obtained in these analyses are also relative.

Not only do analyses with different analysis units answer different questions, different patterns of results across levels of analysis reveal different types of information about teacher behaviors. There were three different patterns of significant results which emerged from the multiple analyses of the 18 teacher behaviors included in this study. Each of these patterns will be discussed below.

For two variables, "students answer teacher questions correctly" and "students fail to respond to teacher questions," significant results were obtained at each level of analysis--individual-within-reading-group, reading-group-within-class, and class. Because these relationships were

consistent across all three levels of analysis, they can be interpreted in a very straightforward manner. Within the range of responses observed in these data, the greater the amount of a behavior, the better (or worse) a student achieved. Accordingly, one would expect a significant relationship both within classes and between classes. For example, classes with a higher percentage of correct answers would achieve more than classes with a lower percentage. Similarly, within classes, students or reading groups with a higher percentage of correct answers would achieve more than students or groups with a lower percentage of correct answers.

The second pattern to be discussed involves nine teacher behavior variables significantly related to achievement at only the individual-within-reading-group level of analysis. These variables include:

student calling out answers to teacher questions,

teacher asking reading questions,

teacher asking nonreading questions,

student responding incorrectly to teacher questions,

student responding to teacher questions with "don't know",

teacher praising student responses,

teacher criticizing student responses,

teacher terminating a student's opportunity to respond, and

teacher sustaining a student's response opportunity.

The presence of a relationship only at the individual-within-reading-group level of analysis suggests that within-group comparisons are responsible for the significant relationship. In other words, the absolute amount of a teacher behavior a student receives is not related to achievement (which may be why the relationship does not show up at the class level of analysis), but

the relative standing of the student within his reading group is related to achievement. For example, teacher criticism of student responses was significantly negatively related to achievement at the student-within-reading-group level of analysis but was unrelated to achievement at the other two analysis levels. This indicates that some type of group context effect may be responsible for the significant relationship at the individual level. In other words, it does not matter how much criticism a student recieves but rather how much he receives relative to his peers in the same reading group. Possibly, students who receive more criticism develop negative feelings, as they know their performance does not compare favorably with that of their peers.

The third pattern involves one teacher behavior variable, "teacher selects a nonvolunteer to respond," which was significant at the class and reading-group-within-class levels of analysis but was not significant at the student-within-reading-group level of analysis. In this case, it was hypothesized that the nature of the variable was responsible for the absence of a relationship at the student-within-reading-group level of analysis. Use of nonvolunteer selection with a reading group was likely to have a similar effect on all students in the group, regardless of who was selected as a nonvolunteer. In other words, nonvolunteer selection is a process by which any student in the reading group could be called on at any time. Hence, all students within a reading group in which the teacher was using nonvolunteer selection would be similarly affected.

One final issue to emerge from these analyses concerns within class variances. The variance among reading group means within classes was relatively small, at least for certain teacher variables such as the teacher terminating or sustaining a students response opportunity. This is an

interesting phenomenon, suggesting that teachers tended to treat reading groups within their classes similarly with respect to these variables, even though the reading groups themselves varied widely in achievement. However, the variance for students within reading groups was considerably greater, indicating that teachers were responding differentially to students within the reading groups. An important area for additional research would be to investigate on what basis, if not achievement, teachers differentiate among students. And, if teachers do differentiate, when such differential treatment of students within a group is appropriate.

In conclusion, both within class and class mean analyses are highly recommended in teacher effectiveness research. Relationships which occur at the class mean level of analysis may not occur within classes and vice versa. Therefore, a combination of both types of analyses will provide a more comprehensive picture of the relationship of student-teacher interactions occuring in the classroom to student achievement.

Summary Table of Results Across All Levels of Analysis

Item	Class			Reading Group*			Student**		
	$\bar{x}$	SD	Results	$\bar{x}$	SD	Results	$\bar{x}$	SD	Results
1. Total initial selections that were ordered	.54	.29	ns	.53	.07	ns	.53	.07	ns
2. Total initial selections that were preselections	.08	.10	ns	.09	.02	ns	.09	.03	ns
3. Total initial selections that were nonvolunteer	.18	.12	.04 (Interaction)	.18	.03	$<.05$(-)	.19	.06	ns
4. Total initial selections that were volunteer	.14	.11	ns	.14	.03	ns	.13	.05	ns
5. Total initial selections that were call outs	.06	.04	ns	.06	.02	ns	.06	.05	$<.01$(-)
6. Total R.O.s that were reading questions	.69	.14	ns	.68	.06	ns	.69	.06	.01(-)
7. Total R.O.s that were nonreading questions	.30	.14	ns	.30	.06	ns	.30	.06	.03(+)
8. Total R.O.s with correct answers	.71	.06	.02(+)	.72	.06	.02(+)	.72	.10	$<.001$(+)
9. Total R.O.s with incorrect answers	.22	.05	ns	.21	.05	ns	.21	.08	$<.001$(-)
10. Total R.O.s with don't know answers	.02	.01	ns	.01	.01	ns	.01	.01	$<.001$(-)
11. Total R.O.s with no response answers	.12	.04	$<.01$(-)	.12	.03	.04(-)	.12	.07	$<.001$(-)

(continued)

	Class			Reading Group			Student		
	$\bar{x}$	SD	Results	$\bar{x}$	SD	Results	$\bar{x}$	SD	Results
12. Correct answers with no feedback	.08	.08	ns	.08	.04	ns	.07	.03	ns
13. Correct answers with emphasis answers	.37	.12	ns	.36	.07	ns	.37	.08	ns
14. Correct answers with praise feedback	.10	.06	ns	.10	.05	ns	.10	.05	$<.01$ (Interaction)
15. Correct answers with new question feedback	.19	.04	.03(+)	.19	.06	ns	.18	.06	ns
16. Incorrect answers with criticism feedback	.02	.01	ns	.02	.02	ns	.02	.03	.03(-)
17. Total incorrect answers with terminating feedback	.49	.14	ns	.50	.06	ns	.49	.15	$<.001$(-)
18. Total incorrect answers with sustaining feedback	.48	.13	ns	.46	.06	ns	.48	.14	$<.01$(+)

*For reading group results, means were calculated for each reading group within a class and then averaged across classes. Standard deviations of reading group means were also computed within each class and then averaged across classes.

**For individual student results, scores were calculated for each student and then averaged across reading groups and classes. Standard deviations of student scores were also computed within each reading group and then averaged across reading groups and classes.

Reference Notes

1. Anderson, L., Evertson, C., and Brophy, J. The First-grade Reading Group Study: Technical Report of Experimental Effects and Process-Outcome Relationships (R&D Report No. 4070). Austin, Texas: Research and Development Center for Teacher Education, 1978.

2. Cronbach, L. Research on classrooms and schools: Formulation of questions, design, and analysis. Occasional Paper of the Stanford Evaluation Consortium, Stanford University, 1976.